Some convergence results for the Newton-GMRES algorithm

Rémi Choquet, Jocelyne Erhel

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Some convergence results for the Newton-GMRES algorithm

Rémi Choquet, Jocelyne Erhel

N° 2065

Septembre 1993

PROGRAMME 6

Calcul scientifique,
modélisation
et logiciels numériques

Some convergence results for the Newton-GMRES algorithm

Rémi Choquet *, Jocelyne Erhel **

Programme 6 — Calcul scientifique, modélisation et logiciel numérique
Projet ALADIN

Rapport de recherche n ° 2065 — Septembre 1993 — 19 pages

Abstract: In this paper, we consider both local and global convergence of the Newton algorithm to solve nonlinear problems when GMRES is used to invert the Jacobian at each Newton iteration. Under weak assumptions, we give a sufficient condition for an inexact solution of GMRES to be a descent direction in order to apply a backtracking technique. Moreover, we extend this result to a finite difference scheme considering also the use of preconditioners. Then we show the impact of the condition number of the Jacobian on the local convergence of the Newton-GMRES algorithm.

Key-words: convergence, descent direction, finite difference, GMRES, Newton, preconditioning.

(Résumé : tsvp)

*choquet@irisa.fr
**erhel@irisa.fr

Unité de recherche INRIA Rennes
IRISA, Campus universitaire de Beaulieu, 35042 RENNES Cedex (France)
Téléphone : (33) 99 84 71 00 – Télécopie : (33) 99 38 38 32

Quelques résultats de convergence pour l'algorithme de Newton-GMRES

Résumé : Dans cet article, nous étudions la convergence de l'algorithme de Newton appliqué à un problème non linéaire, lorsqu'à chaque itération GMRES résout le système linéaire de manière approchée. Une condition suffisante sur la solution calculée par GMRES est donnée pour qu'elle soit une direction de descente afin d'utiliser une technique de «backtracking». Puis, nous étendons ce résultat au cas où chaque produit matrice vecteur est approchée par un schéma aux différences finies, en considérant aussi l'utilisation d'un préconditionnement. Finalement la convergence locale de Newton-GMRES est démontrée dans ce contexte en soulignant l'importance d'un bon préconditionnement.

Mots-clé : convergence, direction de descente, différences finies, GMRES, Newton, préconditionnement.

1 Introduction

The resolution of many physical equations leads to a nonlinear problem. After discretisation of the space considered, such a problem can be written,

$$F(u) = 0 \text{ with } F : \mathbb{R}^n \to \mathbb{R}^n, \tag{1}$$

where $u \in \mathbb{R}^n$ is the unknown. One way to solve (1) is to linearize F but this approach represents only the first step of the well-known Newton method and gives quite poor results in terms of accuracy when F is non-smooth.
Although the Newton method converges only locally, it is commonly used for its good properties of convergence and accuracy.
Furthermore, techniques such as linesearch backtracking strategy improve the global convergence [DS83]. Considering for example the resolution of the Navier-Stokes equations by Newton at each time step, each step of this algorithm involves a large sparse non symmetric matrix to inverse. Due to the structure of this Jacobian, few solvers are efficient on parallel computers. A commonly used algorithm is the so-called GMRES [SS86] (Generalized Minimum RESidual) which gives good results for this kind of matrices. The composition of Newton, GMRES and linesearch backtracking is called non-linear GMRES algorithm and can be summarized by the following algorithm. In the following, we denote by J the jacobian of F.

Algorithm (Newton-GMRES)

u_0 given, $i = -1$
REPEAT
  $i = i + 1$
  Solve $J(u_i)\delta(i) = -F(u_i)$ by GMRES
  $u_{i+1} = u_i + \alpha_i \delta(i)$
UNTIL *convergence*

where α_i is computed by linesearch backtracking to decrease $f(u) = \frac{1}{2}F^t(u)F(u)$.

We know that this algorithm is efficient if the $\delta(i)$ evaluated by GMRES is a descent direction at u_i, i.e

$$F^t(u_i)J(u_i)\delta(i) < 0.$$

In this case, there exists an α satisfying

$$f(u_i + \alpha\delta(i)) < f(u_i).$$

Brown and Saad [BS90] show this result for GMRES without restarting. Brown [Bro87] extends this result within inexact Newton framework when a finite difference scheme is used to approximate Jv, i.e

$$J(u).v \approx \frac{F(u+\sigma v) - F(u)}{\sigma}$$

In this paper, we give an extension of these results considering firstly restarting procedures in section 3, then finite difference scheme with restarting in subsection 4.1 and finally preconditioning in subsection 4.2.

Another subject of interest is to measure the effect of inexact linear solutions over the local convergence of Newton. Under mild assumptions, Brown [Bro87] has considered the local convergence of Newton-GMRES algorithm with a finite difference scheme. In Section 5, we show the same results using the background of section 4. In particular, we study the impact of the condition number of J over the local convergence.

2 GMRES

In this section, we present briefly the GMRES algorithm, introducing some notations we'll need later.

Considering the system

$$J\delta = -F, \tag{2}$$

where $J = J(u)$, $F = F(u)$ and J is non-singular. The principle of GMRES is to minimize $\| - J(\delta_0 + z) - F\|_2$ with z in the Krylov subspace $K(J, r_0, k)$ defined next.

Definition 2.1 *The Krylov subspace $K(J, r_0, k)$ associated to the vector $r_0 = -F - J\delta_0$ and the matrix J is the subspace spanned by $\langle r_0, Jr_0, \ldots, J^{k-1}r_0\rangle$ where δ_0 denotes a first estimation of δ.*

Thus, (2) is replaced by the following minimization problem

$$\min_{z \in K(J, r_0, k)} \|r_0 - Jz\|_2. \tag{3}$$

The Krylov subspace is generated by the Arnoldi process leading to the following GMRES algorithm.

Algorithm (full-GMRES)

given $\varepsilon > 0$ the tolerance of the stopping criterion and $\delta_0 \in \mathbb{R}^n$ the initial guess,

$(s1)$. $r_0 = -F - J\delta_0$, $\beta = \|r_0\|_2$, $v_1 = \dfrac{r_0}{\beta}$, $k = 0$

$(s2)$. REPEAT

$k = k + 1$

(a) $w_{k+1} = Jv_k - \displaystyle\sum_{m=1}^{k} h_{m,k} v_m$

with $h_{m,k} = (Jv_k, v_m)$ $\qquad (m = 1, \ldots, k)$

$h_{k+1,k} = \|w_{k+1}\|_2$

$v_{k+1} = \dfrac{w_{k+1}}{h_{k+1,k}}$

(b) compute $\rho_k = \min\limits_{y \in \mathbb{R}^k} \|\beta e_1 - \overline{H}_k y\|_2$

UNTIL $\rho_k \leq \varepsilon$

$(s3)$. compute y_k, $z_k = V_k y_k$ and $\delta_k = \delta_0 + z_k$

where

$$(\overline{H}_k)_{m,l} = \begin{cases} h_{m,l} & 1 \leq m \leq k+1, m \leq l \leq k \\ 0 & \text{otherwise} \end{cases}$$

$V_k = (v_1, \ldots, v_k)$,

and

$$\|\beta e_1 - \overline{H}_k y_k\|_2 = \min_{y \in \mathbb{R}^k} \|\beta e_1 - \overline{H}_k y\|_2.$$

The Gram-Schmidt process generates an orthogonal basis such that

$$JV_k = V_{k+1} \overline{H}_k, \tag{4}$$

implying

$$\min_{z \in K(J, r_0, k)} \|r_0 - Jz\|_2 = \min_{y \in \mathbb{R}^k} \|\beta e_1 - \overline{H}_k y\|_2, \qquad \beta = \|r_0\|_2. \tag{5}$$

We can easily check that the three following properties ($k \geq 1$) are equivalent.

(i)- $K(J, r_0, k) = K(J, r_0, k-1)$

(ii)- $\rho_k = 0$

(iii)- $J\delta_k = -F$

Note that $\rho_k = \|r_k\|_2$ with $r_k = -F - J\delta_k = r_0 - JV_k y_k$.

Property (ii) gives us an easy criterion to evaluate the state of convergence of this algorithm.

Under the assumption that $\delta_0 = 0$, Brown and Saad [BS90] proved that δ_k $(k \geq 1)$, is a descent direction for F. Let us recall here the result presented in [BS90].

Proposition 2.1 *Given an updated solution u of the Newton-GMRES algorithm presented above, assume that $J = J(u)$ is non-singular and that $\delta_0 = 0$. If $\delta_k \neq 0$ then δ_k $(k \geq 1)$, is a descent direction for f at u.*

$$Furthemore \quad F^t J\delta_k = -F^t F + \rho_k^2 \tag{6}$$

$$and \quad \rho_k = \|F + J\delta_k\|_2 < \|F\|_2. \tag{7}$$

However, the assumption $\delta_0 = 0$ is not satisfied when using restarting or an initial estimation of the solution. We give now a sufficient condition for any δ to be a descent direction. This proposition allows us to extend easily proposition 2.1 to GMRES with restarting, with inexact Jacobian and with preconditioning.

3 Restarting

Restarting procedures were introduced for GMRES to limit the maximum Krylov subspace and reduce both complexity and memory storage. Using previous notations, they consist in restarting GMRES algorithm with the last solution δ_k. As they introduce another loop over GMRES, we need new notations.

We denote by

- GMRES($\delta^{(j-1)}, k_j$)=GMRES with k_j iterations and with initial guess $\delta^{(j-1)}$.

- $K(J, r_{j-1}, k_j)$ the Krylov subspace spanned by the vectors $(r_{j-1}, Jr_{j-1}, \ldots, J^{k_j-1}r_{j-1})$ at step j within the loop of the restarting procedure.
- $\delta^{(0)}$ given, then $\delta^{(j)} = \delta^{(j-1)} + z_{k_j}^{(j)}$, $j \geq 1$ is the direction given by GMRES after j restarts, $z_{k_j}^{(j)}$ is computed within the restart j and is of the form $V_{k_j} y_{k_j}$, corresponding to the Krylov subspace $K(J, r_{j-1}, k_j)$.
- $r^{(j)} = -F - J\delta^{(j)}$, $j \geq 0$ is the residual after j restarts.

GMRES with restarting can be described easily with the following algorithm.

Algorithm (restart-GMRES)

$\delta^{(0)}$ given, $j = 0$
REPEAT
  $j = j + 1$
  Solve $J\delta^{(j)} = -F$ by GMRES$(\delta^{(j-1)}, k_j)$
UNTIL $\|r^{(j)}\|_2 < \varepsilon$.

We now give a sufficient condition for $\delta^{(j)}$ $(j \geq 1)$, to be a descent direction.

Proposition 3.1 *Let $\delta \in \mathbb{R}^n$ and $r = -F - J\delta$, δ is a descent direction as soon as $\|r\|_2 < \|F\|_2$.*

Proof. We have, $F^t J\delta = -F^t F - F^t r$.
Furthermore, $|F^t r| \leq \|F\|_2 \|r\|_2$,
thus $F^t J\delta < 0$ as soon as $\|r\|_2 < \|F\|_2$. □

In order to apply Proposition 3.1, we want to get an inequality on the last residual. We first prove a result on the sequence of residuals from which we can easily deduce the final statement.

Proposition 3.2 *The sequence of the restart-GMRES residuals decreases with a restarting procedure. Furthermore, if J is non-singular then the decrease is strict as soon as $z_{k_j}^{(j)} \neq 0$ $(j \geq 1)$.*

Proof. After each restart, the following residual norm is bounded by the previous one as,

$$\|r^{(j+1)}\|_2 = \min_{z \in K(J, r_j, k_j)} \|r^{(j)} - Jz\|_2 \leq \|r^{(j)}\|_2.$$

If J is non-singular then the solution of (3) is unique, so that $z_{k_j}^{(j)} \neq 0$ implies $\|r^{(j+1)}\|_2 < \|r^{(j)}\|_2$. □

Proposition 3.3 *The approximate solution $\delta^{(j)}$ $(j \geq 1)$ of GMRES with restarting and with $\delta^{(0)} = 0$ is a descent direction if $\delta^{(j)} \neq 0$ and J is non singular.*

Proof. We have $r^{(0)} = -F$ as $\delta^{(0)} = 0$.
Proposition 3.2 gives $\|r^{(j)}\|_2 < \|F\|_2, \quad (j > 0)$.
Thus by Proposition 3.1, $\delta^{(j)}$ is a descent direction. □

Proposition 3.1 cannot be applied when a finite difference scheme is used to approximate Jv, since the considered residual is an approximation of the exact residual. However, we now state a sufficient condition on the norm of the residual to keep a descent direction, again in the case where $\delta \neq 0$.

4 Inexact Jacobian

Sometimes, Jacobians are very difficult to estimate and/or need heavy memory storage. In these cases, as GMRES requires merely the product of the Jacobian by a vector, we use a finite difference scheme to estimate Jv. With this approach, the non-linear GMRES algorithm belongs to the class of Inexact Newton's methods.

Following Brown [Bro87] ideas, inexact-GMRES algorithm can be described as a GMRES algorithm with a perturbed system. By this way and with the same approach as in Proposition 3.1, we give a sufficient condition for the inexact-GMRES solution to be a descent direction. Inside inexact-GMRES algorithm, we consider also the use of a right preconditioner. In general, the use of a preconditioner involves a new linear system to be solve. Due to the cost, we solve it approximately. In subsection 4.2 we take into account the errors introduced by this approximation to show a result similar to Proposition 3.1.

4.1 Algorithm

Let us first define the GMRES algorithm with the use of a finite difference scheme, which we shall call inexact-GMRES. In order to simplify notations, we shall not consider explicitly restarting procedures. But, the further results can be extended easily to this case. We overline each variable of GMRES by a ($\tilde{}$) to exhibit the difference between the two algorithms.

Algorithm (inexact-GMRES)

given $\varepsilon > 0$ the tolerance of the stopping criterion and $\delta_0 \in \mathbb{R}^n$ the initial guess,

$(s1)$. $\tilde{r}_0 = -F - q_1$, $q_1 = \dfrac{F(u+\sigma_0\delta_0) - F(u)}{\sigma_0}$

$\tilde{\beta} = \|\tilde{r}_0\|_2$, $\tilde{v}_1 = \dfrac{\tilde{r}_0}{\tilde{\beta}}$

$k = 0$

$(s2)$. REPEAT

$k = k + 1$

(a) $q_{k+1} = \dfrac{F(u + \underset{k}{\sigma_k}\tilde{v}_k) - F(u)}{\sigma_k}$

$\tilde{w}_{k+1} = q_{k+1} - \displaystyle\sum_{m=1}^{k} \tilde{h}_{m,k}\tilde{v}_m$

with $\tilde{h}_{m,k} = (q_{k+1}, \tilde{v}_m) \qquad (m = 1, \ldots, k)$

$\tilde{h}_{k+1,k} = \|\tilde{w}_{k+1}\|_2$

$\tilde{v}_{k+1} = \dfrac{\tilde{w}_{k+1}}{\tilde{h}_{k+1,k}}$

(b) compute $\tilde{\rho}_k = \min\limits_{y \in \mathbb{R}^k} \|\tilde{\beta}e_1 - \tilde{\overline{H}}_k y\|_2$

UNTIL $\tilde{\rho}_k \leq \varepsilon$

$(s3)$. $K = k$

compute $\tilde{y}_K$, $\tilde{z}_K = \tilde{V}_K \tilde{y}_K$ and $\tilde{\delta}_K = \delta_0 + \tilde{z}_K$

where

$$(\tilde{\overline{H}}_k)_{m,l} = \begin{cases} \tilde{h}_{m,l} & 1 \leq m \leq k+1, m \leq l \leq k \\ 0 & \text{otherwise} \end{cases}, \tag{8}$$

$$\tilde{V}_k = (\tilde{v}_1, \ldots, \tilde{v}_k), \tag{9}$$

and

$$\|\tilde{\beta} e_1 - \tilde{\overline{H}}_k \tilde{y}_k\|_2 = \min_{y \in \mathbb{R}^k} \|\tilde{\beta} e_1 - \tilde{\overline{H}}_k y\|_2. \tag{10}$$

As Brown [Bro87], let us define an error matrix $E_k = \Sigma_k \tilde{V}_k^{\,t}$ with $\Sigma_k = (\epsilon_1, \ldots, \epsilon_k)$ and $\epsilon_k = q_{k+1} - J\tilde{v}_k \quad (k = 1, \ldots, K)$.

In the sequel, we will denote E_K as E.

According to the previous definitions and since $Ev_k = \epsilon_k$, we have the following property,

$$(J + E)\tilde{v}_k = q_{k+1} \text{ for } k = 1, \ldots, K. \tag{11}$$

Now, we introduce also an error in $\tilde{r}_0$, due to an initial non null δ_0; hence, we define $\epsilon_0 = q_1 - J\delta_0$.
Let us denote

$$\begin{cases} \tilde{F} = F + \epsilon_0 - E\delta_0 \\ \tilde{J} = J + E \end{cases}$$

We show that the inexact-GMRES algorithm is equivalent to GMRES algorithm (section 2) where we set $J = \tilde{J}$ and $F = \tilde{F}$.

This result is based on the two next lemmas and follows the lines of Brown's proof except that here $\delta_0 \neq 0$ involves a new right-hand side in the equivalent GMRES procedure.

Lemma 4.1 *$\overline{H}_k = \tilde{\overline{H}}_k$ and $V_{k+1} = \tilde{V}_{k+1}$ for all $k = 1, \ldots, K$.*

Proof. We use the notations with no tilde for the GMRES version.
We will make this proof by induction.
With $F = \tilde{F}$ and $J = \tilde{J}$ in GMRES,

$$r_0 = -(F + \epsilon_0 - E\delta_0) - (J + E)\delta_0 = -F - (\epsilon_0 + J\delta_0) = -F - q_1 = \tilde{r}_0.$$

So, $v_1 = \tilde{v}_1$.

Starting step $(s2)$ of both algorithms with the same initial condition, we then have $\overline{H}_1 = \tilde{\overline{H}}_1$ and $V_2 = \tilde{V}_2$ since

$$\begin{aligned} & h_{1,1} = ((J+E)v_1, v_1) = (q_2, \tilde{v}_1) = \tilde{h}_{1,1}, \\ & \tilde{w}_2 = q_2 - \tilde{h}_{1,1}\tilde{v}_1 = (J+E)v_1 - h_{1,1}v_1 = w_2, \\ \text{hence} \quad & v_2 = \tilde{v}_2 \text{ and } h_{2,1} = \tilde{h}_{2,1}. \end{aligned}$$

Now, assuming that $\overline{H}_{k-1} = \tilde{\overline{H}}_{k-1}$ and $V_k = \tilde{V}_k$, we obtain at the next step $\overline{H}_k = \tilde{\overline{H}}_k$ and $V_{k+1} = \tilde{V}_{k+1}$, since for $m = 1, \ldots, k$

$$\begin{aligned} & h_{m,k} = ((J+E)v_k, v_m) = (q_{k+1}, \tilde{v}_m) = \tilde{h}_{m,k}, \\ & \tilde{w}_{k+1} = q_{k+1} - \sum_{m=1}^{k} \tilde{h}_{m,k}\tilde{v}_m = (J+E)v_k - \sum_{m=1}^{k} h_{m,k}v_m = w_{k+1}, \\ \text{hence} \quad & v_{k+1} = \tilde{v}_{k+1} \text{ and } h_{k+1,k} = \tilde{h}_{k+1,k}. \end{aligned}$$

□

We derive the following property.

Lemma 4.2 *Assuming that $J+E$ is non-singular then for all $k \leq K$,*

$$\begin{aligned} & \rho_k = \tilde{\rho}_k = \min_{y \in \mathbb{R}^k} \|r_0 - (J+E_k)V_k y\|_2 \\ & \textit{and } \delta_k = \tilde{\delta}_k. \end{aligned}$$

Proof.

$$\begin{aligned} \tilde{\rho}_k &= \min_{y \in \mathbb{R}^k} \|\tilde{\beta} e_1 - \tilde{\overline{H}}_k y\|_2 \\ &= \min_{y \in \mathbb{R}^k} \|\beta e_1 - \overline{H}_k y\|_2 \\ &= \rho_k \end{aligned}$$

Thus, $y_k = \tilde{y}_k$ as $J+E$ is non-singular and $\delta_k = \tilde{\delta}_k$. □

Let us define the residual computed by the inexact-GMRES algorithm. By Lemma 4.2, we have $\tilde{\rho}_k = \rho_k = \|r_k\|_2$ where r_k is the residual computed by the equivalent GMRES, i.e

$$r_k = r_0 - (J+E)V_k y_k \ ;$$

so let us define

$$\tilde{r}_k = \tilde{r}_0 - (J+E)\tilde{V}_k \tilde{y}_k.$$

Now we can derive a sufficient condition to get a descent direction,

Proposition 4.1 *Assuming that $J+E$ is non-singular then $\tilde{\delta}_k,\ k \leq K$ is a descent direction for J if,*

$$\tilde{\rho}_k < \|F\|_2 - \|\epsilon_0\|_2 - \|\Sigma_k\|_2 \|\tilde{y}_k\|_2, \tag{12}$$

where $\epsilon_0 = q_1 - J\delta_0$, $\Sigma_k = (\epsilon_1, \ldots, \epsilon_k)$ and $\epsilon_k = q_{k+1} - J\tilde{v}_k$.

Proof.

$$\begin{aligned} J\tilde{\delta}_k &= J\delta_0 + J\tilde{V}_k\tilde{y}_k \\ &= -(F + \epsilon_0 + \tilde{r}_0) + J\tilde{V}_k\tilde{y}_k) \\ &= -(F+\epsilon_0) - \tilde{r}_k - E\tilde{V}_k\tilde{y}_k \\ F^t J\tilde{\delta}_k &\leq \|F\|_2(\|\tilde{r}_k\|_2 + \|\epsilon_0\|_2 + \|\Sigma_k\|_2\|\tilde{y}_k\|_2 - \|F\|_2) \end{aligned}$$

The proposition follows. □

Thus, bounds for $\|\epsilon_k\|,\ k = 0, \ldots, K$ and (12) give a criterion over $\tilde{\delta}_k,\ k \leq K$ to be a descent direction.

4.2 Inexact GMRES with a right preconditioner

In this part, we present the inexact-GMRES algorithm with a right preconditioner B and we extend the results presented in subsection 4.1.
System (2) becomes

$$\begin{cases} JB^{-1}z = -F \\ B\delta = z \end{cases} \tag{13}$$

so that each iteration of inexact-GMRES algorithm involves a linear resolution

$$Bx_k = \tilde{v}_k. \tag{14}$$

where the vectors $\tilde{v}_k \quad (k = 1, \ldots, K)$ define the basis of the Krylov subspace $K(JB^{-1}, \tilde{r}_0, K)$ constructed by the Arnoldi process. In the following, we consider the practical case where (14) is solved approximatively. We first recall the preconditioned inexact GMRES algorithm, where we denote by τ_k the error on the solution in (14).

Algorithm (pre-inexact-GMRES)

given $\varepsilon > 0$ the tolerance of the stopping criterion and $\delta_0 \in \mathbb{R}^n$ the initial guess,

$(s1)$. $\tilde{r}_0 = -F - q_1$, $q_1 = \dfrac{F(u+\sigma_0\delta_0) - F(u)}{\sigma_0}$

$\tilde{\beta} = \|\tilde{r}_0\|_2$, $\tilde{v}_1 = \dfrac{\tilde{r}_0}{\tilde{\beta}}$

$k = 0$

$(s2)$. REPEAT

$k = k+1$

(a) $\tilde{x}_k$ an approximate solution of $Bx_k = \tilde{v}_k$
with $\tilde{x}_k = B^{-1}\tilde{v}_k + \tau_k$

(b) $q_{k+1} = \dfrac{F(u+\sigma_k\tilde{x}_k) - F(u)}{\sigma_k}$

$\tilde{w}_{k+1} = q_{k+1} - \sum_{m=1}^{k} \tilde{h}_{m,k}\tilde{v}_m$

with $\tilde{h}_{m,k} = (q_{k+1}, \tilde{v}_m) \qquad (m = 1, \ldots, k)$

$\tilde{h}_{k+1,k} = \|\tilde{w}_{k+1}\|_2$

$\tilde{v}_{k+1} = \dfrac{\tilde{w}_{k+1}}{\tilde{h}_{k+1,k}}$

(c) compute $\tilde{\rho}_k = \min_{y\in\mathbb{R}^k} \|\tilde{\beta}e_1 - \tilde{\overline{H}}_k y\|_2$

UNTIL $\tilde{\rho}_k \leq \varepsilon$

$(s3)$. $K = k$

compute $\tilde{y}_K$

$(s4)$. $\tilde{\delta}_K = \delta_0 + \tilde{X}_K\tilde{y}_K$

where $\tilde{\overline{H}}_k$,$\tilde{V}_k$ and $\tilde{y}_k$ are defined by (8,9,10) and $\tilde{X}_k = (\tilde{x}_1, \ldots, \tilde{x}_k)$.

Let $T_k = (\tau_1, \ldots, \tau_k)$.
At step $(s4)$ of the previous algorithm, an exact update of the solution would be $\tilde{\delta}_K = \delta_0 + B^{-1}\tilde{V}_K\tilde{y}_K$. In general, such an update is very costly. But, as $B^{-1}\tilde{V}_K\tilde{y}_K = (\tilde{X}_K - T_K)\tilde{y}_K$, we have replaced it by the following approximate formula,

$$(s4). \quad \tilde{\delta}_K = \delta_0 + \tilde{X}_K\tilde{y}_K.$$

Now, let us define

$$\begin{cases} \epsilon_0 = q_1 - J\delta_0 \\ \epsilon_k = q_{k+1} - J\tilde{x}_k \qquad k = 1, \ldots, K \end{cases}$$

and $E_k = (JT_k + \Sigma_k)\tilde{V_k}^t$ with $\Sigma_k = (\epsilon_1, \ldots, \epsilon_k)$.

With the previous definitions and $E = E_K$, we have the following property,

$$(JB^{-1} + E)\tilde{v}_k = q_{k+1} \text{ for } k = 1, \ldots, K. \tag{15}$$

Proof.

$$\begin{aligned} & E\tilde{v}_k = JT_K\tilde{V}_K^t\tilde{v}_k + \Sigma_K\tilde{V}_K^t\tilde{v}_k = J\tau_k + \epsilon_k \\ \text{so} \quad & (JB^{-1} + E)\tilde{v}_k = J(B^{-1}\tilde{v}_k + \tau_k) + \epsilon_k = q_{k+1}. \end{aligned}$$

□

Let us denote,

$$\begin{cases} \tilde{F} = F + \epsilon_0 - E\delta_0 \\ \tilde{J} = JB^{-1} + E \end{cases}$$

If we take $J = \tilde{J}$ and $F = \tilde{F}$ in GMRES (section 2) then steps 1,2,3 of pre-inexact-GMRES algorithm are equivalent to steps 1,2,3 of GMRES algorithm, as the lemmas 4.1 and 4.2 remain true. Proofs are the same as those of subsection 4.1. Let us define as before the residual,

$$\tilde{r}_k = \tilde{r}_0 - (JB^{-1} + E)\tilde{V}_k\tilde{y}_k.$$

By lemma 4.2, we have $\tilde{\rho}_k = \|\tilde{r}_k\|_2$. Furthermore, we have this result.

Proposition 4.2 *Assuming that $JB^{-1} + E$ is non-singular then $\tilde{\delta}_k$, $k \leq K$ is a descent direction for J if*

$$\tilde{\rho}_k < \|F\|_2 - \|\epsilon_0\|_2 - \|\Sigma_k\|_2\|\tilde{y}_k\|_2. \tag{16}$$

Proof.

$$\begin{aligned} \tilde{r}_k &= \tilde{r}_0 - (JB^{-1} + E)\tilde{V}_k\tilde{y}_k \\ &= -(F + \epsilon_0 + J\delta_0) - J(\tilde{X}_k - T_k)y_k - EV_k y_k \\ &= -(F + \epsilon_0) - J\delta_0 - J\tilde{X}_k y_k - \Sigma_k y_k \end{aligned}$$

Thus, $J\tilde{\delta}_k = J\delta_0 + J\tilde{X}_k y_k = -F - \epsilon_0 - \Sigma_k y_k - \tilde{r}_k$. The proposition follows. □

The criterion (16) can be used if bounds for Σ_k and ϵ_0 are known explicitely. In the next subsection, we give such bounds but we'll see that they depend on the Lipschitz constant of J which is in general not easy to estimate.

Remark 4.1 *The criterion (16) is independant of τ_k. With an exact step (s4) of the pre-inexact-GMRES algorithm, (16) would be*

$$\tilde{\rho}_k < \|F\|_2 - \|\epsilon_0\|_2 - \|\Sigma_k\|_2\|\tilde{y}_k\|_2 - \|JT_k\|_2\|\tilde{y}_k\|_2.$$

4.3 Error estimates

In this subsection, we estimate the errors due to the finite difference process, by deriving bounds from the following lemma.

Lemma 4.3 *Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be continuously differentiable in an open convex set $D \subset \mathbb{R}^n$, $u \in D$, and let J be γ Lipschitz continuous in the neighbourhood D. Then, for any $u + p \in D$,*

$$\|F(u+p) - F(u) - J(u)p\| \leq \frac{\gamma}{2}\|p\|^2. \tag{17}$$

A proof is given in [DS83].

We deduce readily bounds on the errors ϵ_k.

Lemma 4.4 *Let $B \in \mathbb{R}^{n \times n}$ non-singular and $r = \max_{k=1,\ldots,K} \|\sigma_k(B^{-1}v_k + \tau_k)\|_2$. We assume that J is γ lipschitz continuous on D an open convex set such that $N(u,r) \subset D$ then,*

$$k \geq 1 \quad \|\epsilon_k\|_2 \leq \frac{\sigma_k}{2}\gamma\|B^{-1}v_k + \tau_k\|_2^2 \tag{18}$$

$$\textit{and} \qquad \|\epsilon_0\|_2 \leq \frac{\sigma_0}{2}\gamma\|\delta_0\|_2^2. \tag{19}$$

Proof. We apply lemma 4.3 with $p = \sigma_k(B^{-1}v_k + \tau_k)$ for $k \geq 1$ and $p = \sigma_0\delta_0$ for $k = 0$. □

Lemma 4.4 gives a bound for the matrix norm $\|E\|_2$.

Corollary 4.1

$$\|E\|_2 = \|\Sigma_K\|_2 \leq \frac{\sqrt{K}\gamma}{2} \max_{k=1,K}(\sigma_k\|B^{-1}v_k + \tau_k\|_2^2)$$

Proof.

$$\begin{aligned}
\|\Sigma_K x\|_2 &\leq \sum_{k=1}^{K} \|\epsilon_k x_k\|_2 \\
&\leq \sum_{k=1}^{K} \|\epsilon_k\|_2 |x_k| \\
&\leq \max_{k=1,\ldots,K} \|\epsilon_k\|_2 \|x\|_1 \\
&\leq \sqrt{K} \max_{k=1,\ldots,K} \|\epsilon_k\|_2 \|x\|_2
\end{aligned}$$

Then, we use lemma 4.4 to show the result. □

Remark 4.2 *(i)- If $B = I$ then $\tau_k = 0$ and $\|\epsilon_k\|_2 \leq \frac{\sigma_k}{2}\gamma$.*

(ii)- τ_k can be estimated using the condition number $\chi(B)$, as

$$\frac{\|\tau_k\|}{\|B^{-1}v_k\|} \leq \chi(B)\|v_k - B\tilde{x}_k\|.$$

(iii)- The main difficulty for checking if δ_k is a descent direction is to bound γ, which can be roughly estimated by using a finite difference approximation of the second derivative at u,

$$\frac{F(u+\sigma p) - 2F(u) + F(u-\sigma p)}{\sigma^2}.$$

Up to now, the errors due to the approximation of the Jacobian by a finite difference scheme have been bounded to show their impact on descent directions hence on global convergence. In the next section, we focus on the local convergence of Newton when the linear system is solved approximately. In particular, we show that preconditioning is an important issue to reduce the relative cost of the linear system solving in the Newton process.

5 Newton convergence with inexact solutions of linear systems

We consider the local convergence of the Newton algorithm with a finite difference scheme and an inexact solution of the linear system. In [Bro87], Brown showed the local convergence under a assumption of a sufficient decrease of the exact residual norm. We use the background of section 4.1 to obtain the same result, with an assumption on the norm of the residual computed at each step of the inexact-GMRES algorithm. We show that under some assumptions, Newton converges if the solution of the linear system at each step of Newton is accurate enough. As a linear solver, we take here inexact-GMRES described in subsection 4.1 with a restarting. But another linear system solver can be considered. GMRES algorithm with restarting is equivalent to GMRES with a non null initial estimation of the solution, so that Newton-restart-inexact-GMRES algorithm can be written (recall that $J_i = J(u_i)$ and $F_i = F(u_i)$) in the following framework.

Algorithm (Newton-restart-inexact-GMRES)

u_0 given, $i = -1$
REPEAT (Newton loop)
$i = i + 1$
$(s1)$. Solve $J_i\delta(i) = -F_i$ by (restarted-inexact-GMRES)
with a first estimation $\delta^{(0)}(i)$ and an ending criterion ε_i.
$\tilde{\delta}(i)$ the solution at convergence.
$(s2)$. $u_{i+1} = u_i + \tilde{\delta}(i)$
UNTIL *convergence*.

The results will be given in the l_2 induced matrix norm.
As we consider inexact-GMRES algorithm, we extend some notations of subsection 4.1, adding a new index to the variables E and ϵ_0 noted from now on $E(i)$ and $\epsilon_0(i)$. Recall that bounds for E and ϵ_0 depend linearly from σ_k (subsection 4.3). To simplify, we note $\sigma(i)$ the maximum of σ_k used in stage $(s1)$ of the previous algorithm. We consider the last restart of GMRES where the convergence is obtained and where we note by $\delta_0(i)$ the initial guess which is in fact the solution obtained by the previous restart.
Furthermore, as shown in subsection 4.1, the inexact-GMRES algorithm is equivalent to solve

$$\tilde{J}_i\delta(i) = -\tilde{F}_i \text{ by GMRES (section 2)},$$

where $\tilde{J}_i = J_i + E(i)$ and $\tilde{F}_i = F_i + \epsilon_0(i) - E(i)\delta_0(i)$.
As previously, we note $\tilde{r}_i = -\tilde{F}_i - \tilde{J}_i\tilde{\delta}(i)$ where $\tilde{\delta}(i)$ is the solution at convergence, so that $\|\tilde{r}_i\| \leq \epsilon_i$;
Thus each Newton step is of the form

$$u_{i+1} = u_i - \tilde{J}_i^{-1}\tilde{F}_i - \tilde{J}_i^{-1}\tilde{r}_i.$$

We give now the main result of this section.

Proposition 5.1 *Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be continuously differentiable in an open convex set $D \subset \mathbb{R}^n$. Assume that there exists u_* and $r, \beta > 0$, such that,*

$$\begin{cases} N(u_*, r) \subset D \\ F(u_*) = 0, \quad J(u_*)^{-1} exists \text{ and } \|J(u_*)^{-1}\| \leq \beta \\ J \in Lip_\gamma(N(u_*, r)) \end{cases}$$

Solving each step of Newton by inexact-GMRES with restarting lead to a well-defined and convergent sequence $(u_i)_i$ if the convergence in inexact-GMRES is obtained with a tolerance ϵ_i small enough and if each parameter $\sigma(i)$ in the finite difference is small enough.

Proof. This result is based on the following theorem.

Theorem 5.1 *Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be continuously differentiable in an open convex set $D \subset \mathbb{R}^n$. Assume that there exists u_* and $r, \beta > 0$, such that,*

$$\begin{cases} N(u_*, r) \subset D \\ F(u_*) = 0, \quad J(u_*)^{-1} exists \text{ and } \|J(u_*)^{-1}\| \leq \beta \\ J \in Lip_\gamma(N(u_*, r)) \end{cases}$$

Let a sequence $u_1, u_2, \ldots$ generated by

$$u_{i+1} = u_i - \tilde{J}_i^{-1}\tilde{F}_i + \tilde{J}_i^{-1}\tilde{r}_i \text{ where } \tilde{J}_i = J_i + E(i),$$

with

(A1)- $\|E(i)\| \leq \dfrac{\theta_i}{3\eta}\|F_i\|$,

(A2)- $\|\tilde{F}_i - F_i\| \leq \dfrac{\theta_i}{3}\|F_i\|$,

(A3)- $\|\tilde{r}_i\| \leq \dfrac{\theta_i}{3}\|F_i\|$,

and a real sequence $(\theta_i)_i$ and a scalar $r > \eta > 0$ satisfying

$$\beta\{\gamma\eta + 2M\theta_i\} \leq \frac{1}{2}, \tag{20}$$

with $M = \sup_{u \in N(u_*, \eta)} \|J(u)\|$.
Then for all $u_0 \in N(u_*, \eta)$, the sequence $u_1, u_2, \ldots$ is well defined and converges linearly to u_*. More precisely, we have

$$\|u_{i+1} - u_*\| \leq \frac{1}{2}\|u_i - u_*\|.$$

Proof. We prove the theorem by recurrence, assuming $u_i \in N(u_*, \eta)$, which is true for $i = 0$.

- First, we show that $\tilde{J}_i$ is non-singular. It's based on the following lemma.

  Lemma 5.1 *If A is non-singular and $\|A^{-1}(B - A)\| < 1$, then B is non-singular and* $\|B^{-1}\| \leq \dfrac{\|A^{-1}\|}{1 - \|A^{-1}(B - A)\|}$.

  See [DS83] for the demonstration.

  We have

$$\begin{aligned} \|J(u_*)^{-1}[\tilde{J}_i - J(u_*)]\| &\leq \|J(u_*)^{-1}\|\|[\tilde{J}_i - J_i)] + [J_i - J(u_*)]\| \\ &\leq \beta(\|E(i)\| + \gamma\eta) \\ &\leq \beta(\frac{\theta_i}{3}M + \gamma\eta) \\ &\leq \frac{1}{2}. \end{aligned}$$

  By Lemma 5.1 then $\tilde{J}_i$ is inversible and (20) gives

$$\|\tilde{J}_i^{-1}\| \leq 2\beta. \tag{21}$$

- Therefore u_{i+1} is well defined and

$$\begin{aligned} u_{i+1} - u_* &= u_i - u_* - \tilde{J}_i^{-1}\tilde{F}_i + \tilde{J}_i^{-1}\tilde{r}_i \\ &= \tilde{J}_i^{-1}\{F(u_*) - F_i - J_i(u_* - u_i) - E(i)(u_* - u_i) - (\tilde{F}_i - F_i) + \tilde{r}_i\} \end{aligned}$$

  By Lemma (4.3) $\|F(u_*) - F_i - J_i(u_* - u_i)\| \leq \dfrac{\gamma}{2}\|u_i - u_*\|^2$.

  It should be noted that the inequality above is the quadratic term in the exact Newton procedure. Now we deal with the errors due to the inexact Jacobian and to the approximate solution by GMRES.

  By (A1-3) and using $\|F_i\| \leq M\|u_i - u_*\|$, we get

$$\begin{cases} \|E(i)(u_* - u_i)\| & \leq \|E(i)\|\|u_i - u_*\| \leq \dfrac{\theta_i}{3}M\|u_i - u_*\| \\ \|\tilde{F}_i - F_i\| & \leq \dfrac{\theta_i}{3}\|F_i\| \leq \dfrac{\theta_i}{3}M\|u_i - u_*\| \\ \|\tilde{r}_i\| & \leq \dfrac{\theta_i}{3}\|F_i\| \leq \dfrac{\theta_i}{3}M\|u_i - u_*\| \end{cases}$$

Thus

$$\begin{aligned} \|u_{i+1} - u_*\| &\leq \beta(\gamma\eta + 2\theta_i M)\|u_i - u_*\| \\ &\leq \frac{1}{2}\|u_i - u_*\| \quad \text{by (20).} \end{aligned}$$

□

In order to prove Proposition 5.1, we want to apply Theorem 5.1 hence we have to check the assumptions (A1-3). Fixing θ_i, we define a real sequence $\sigma(i)$ such that (A1) and (A2) are both satisfied. By Lemma 4.4 and Corollary 4.1, we have

$$\begin{cases} \|\tilde{F}_i - F_i\| = \|\epsilon_0(i) - E(i)\delta_0(i)\| \leq \dfrac{\sigma(i)\gamma}{2}\|\delta_0(i)\|(\|\delta_0(i)\| + \sqrt{K}) \\ \|E(i)\| \leq \dfrac{\sigma(i)\gamma}{2}\sqrt{K} \end{cases}$$

We take

$$\sigma(i) = \frac{2\theta_i}{3\gamma}\|F_i\| \min\Big(\frac{1}{\|\delta_0(i)\|(\|\delta_0(i)\| + \sqrt{K})}, \frac{1}{\eta\sqrt{K}}\Big).$$

Then, we choose ε_i the ending criterion of GMRES such that $\varepsilon_i = \dfrac{\theta_i}{3}\|F_i\|$, thus (A3) is satisfied. As (A1-3) are satisfied by (inexact-GMRES), Theorem 5.1 shows us that Newton-restart-inexact-GMRES converges. □

The inequality (20) shows that the local convergence depends on the factor $\beta M\varepsilon_i$. As βM reflects the condition number of the matrix J_i, the linear system has to be solved with more accuracy when the Jacobian is ill-conditioned. Unfortunately, this constraint is quite difficult to obtain when the condition number of J is high. This is another reason to require a good preconditioning to achieve the convergence of Newton at a reasonable cost.

6 Conclusion

In this paper, we have studied the convergence of a Newton-Krylov method where the Jacobian at each Newton iteration is inverted by a GMRES linear solver. For some variants of GMRES, a sufficient condition is given for a solution to be a descent direction, in order to apply a backtracking technique to improve the global convergence of the Newton iterations. This sufficient condition is based on the residual or on the modified residual when finite difference and preconditioning are considered.

The modified residual is far from the exact one if the errors due to finite difference and preconditioning are large. We have obtained bounds on the errors due to the inexact Jacobian which depend on the Lipschitz constant of the Jacobian and naturally on the scalar step chosen in the finite-difference scheme. Since this constant is difficult to estimate, the residual is required in practice to be small enough in order to guarantee a descent direction. However, a good preconditioning is mandatory to achieve this accuracy of convergence at reasonable cost.
Moreover, the residual and the scalar step must both be small enough to ensure a linear local convergence of Newton. Once again, the rate of convergence must be enhanced by a preconditioning to decrease the cost of computation.
Though our results are proved within the GMRES framework, they are actually based on the residual and the ending criterion. Therefore they can be extended to any other matrix-free iterative solver.

References

[Bro87] P.N Brown. A local convergence theory for combined inexact-newton/finite difference projection methods. *SIAM J.Numer.Anal*, 24:407–434, 1987.

[BS90] P.N Brown and Y Saad. Hybrid krylov methods for nonlinear systems of equations. *SIAM J.Sci.Stat.Comput*, 11(3):450–481, Mai 1990.

[DS83] J.E Dennis and R.B Schnabel. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations.* Prentice-Hall series in Computational Mathematics, 1983.

[SS86] Y Saad and H Schultz. Gmres: a generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Statist. Comput.*, 7:856–869, 1986.

Unité de recherche INRIA Lorraine, Technôpole de Nancy-Brabois, Campus scientifique,
615 rue de Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unité de recherche INRIA Rennes, IRISA, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unité de recherche INRIA Rhône-Alpes, 46 avenue Félix Viallet, 38031 GRENOBLE Cedex 1
Unité de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
Unité de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
ISSN 0249-6399